



TITLE:

A domain swapping approach to elucidate differential regiospecific hydroxylation by geraniol and linalool synthases from perilla.

AUTHOR(S):

Sato-Masumoto, Naoko; Ito, Michiho

---

CITATION:

Sato-Masumoto, Naoko ...[et al]. A domain swapping approach to elucidate differential regiospecific hydroxylation by geraniol and linalool synthases from perilla.. *Phytochemistry* 2014, 102: 46-54

ISSUE DATE:

2014-06

URL:

<http://hdl.handle.net/2433/187741>

RIGHT:

© 2014 Elsevier Ltd.; This is not the published version. Please cite only the published version.; この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。

**Title** Domain swapping approach to regiospecific hydroxylation by geraniol and linalool synthases from perilla.

Naoko Sato-Masumoto and Michiho Ito\*

### Abstract

Geraniol and linalool are acyclic monoterpenes found in plant essential oils that have attracted much attention for their commercial use and in pharmaceutical studies. These monoterpenes are synthesized from geranyl diphosphate (GDP) by geraniol and linalool synthases, respectively. Both synthases are very similar at the amino acid level and share the same substrate; however, the position of the GDP to which they introduce hydroxyl groups is different. In this study, we investigated the mechanisms underlying the regiospecific hydroxylation of geraniol and linalool synthases using a domain swapping approach and site-directed mutagenesis in perilla. Sequences of the synthases were divided into ten domains (domains I to IV-4), and each corresponding domain was exchanged between both enzymes. It was shown that different regions were important for the formation of geraniol and linalool, namely, domains IV-1 and -4 for geraniol, and domains III-b, III-d, and IV-4 for linalool. These results suggested that the conformation of carbocation intermediates and their electron localization were seemed to be different between geraniol and linalool synthases. Further, five amino acids in domain IV-4 were indispensable for the formation of geraniol and linalool. According to three-dimensional structural models of the synthases, these five residues seemed to be responsible for the different spatial arrangement of the amino acid at H524 in the case of geraniol synthase, while N526 is the corresponding residue in linalool synthase. Results suggested that the side chains of these five amino acids in combination with several relevant domains localized the positive charge in the carbocation intermediate to determine the position of the introduced hydroxyl group.

### Keywords

*Perilla*; Labiatae; Geraniol synthase; Linalool synthase; Domain swapping; Site-directed mutagenesis; Chimeric enzymes; Regiospecific hydroxylation; Essential oil

## 1. Introduction

Monoterpenes comprise a chemically diverse group of C<sub>10</sub> compounds composed of two isoprene units; they are generally volatile and fragrant and are often found in plant essential oils. In the plant body, monoterpenes play important roles in defense against microbes and insect pathogens and in the attraction of pollinators (Chappell, 1995; Chen et al., 2003; Yamasaki et al., 2007). Many monoterpene compounds are used for perfumery and flavoring of either food or medicine, and quite a few have antibacterial or antitumor properties (Chung et al., 2006; Honda et al. 1984); thus, they are attractive compounds in the field of pharmaceutical sciences.

Perilla, an annual Asian herb, contains essential oil that is composed of unique monoterpene compounds and is designated as a medicine in the Japanese Pharmacopoeia. The oil types of perilla are classified into more than 10 groups according to their principal constituent (Ito et al., 1999). Most of the oil types are monoterpene types and are presumed to be biosynthesized from geranyl diphosphate (GDP, 1) as a starting compound (Fig. 1). This biosynthetic pathway of perilla oil is genetically controlled (Honda, 1996), and the functions of each gene can be determined by cloning the enzymes that catalyze the relevant reaction steps in the biosynthetic pathway.

The first step in the putative biosynthetic pathway of monoterpene type oil components is the dephosphorylation of GDP by either geraniol or limonene synthase (Fig. 1). Linalool (2), which is found in all perilla plants regardless of oil type (Pichersky et al. 1995), is also converted from GDP by linalool synthase. The geraniol and linalool synthases of *Perilla frutescens*, *P. citriodora*, *P. hirtella*, and *P. setoyensis* and the limonene synthases of *P. frutescens* and *P. citriodora* were cloned and shown to be highly similar at the sequence level (Ito and Honda, 2007; Masumoto et al., 2010; Yuba et al., 1996). Among these three monoterpene synthases, geraniol and linalool synthases have highly homologous sequences that are more than 70% identical at the amino acid level (Masumoto et al., 2010) and are considered to convert GDP in a similar manner; GDP is converted to a geranyl cation, and a hydroxyl group derived from water molecules near the active site is added to the intermediate carbocation (Iijima et al., 2004; Jia et al., 1999). As for linalool synthase, it is also conceivable that GDP was changed into linalool via linalyl diphosphate (LDP) and linalyl cation

(Degenhardt et al., 2009). However, detailed reaction intermediate for linalool synthase were unclear. The position of the additional hydroxyl group differs between geraniol (3) and linalool; the former is hydroxylated at C-1, while the latter is hydroxylated at C-3 (Fig. 2). It is of great interest to elucidate the mechanism and identify specific domains or amino acids relevant to the regiospecific hydroxylation of GDP. Since linalool and the other compounds formed via geraniol are found in perilla essential oil, the corresponding synthases for these compounds, namely, linalool and geraniol synthases, are supposed to coexist in perilla, indicating that these two synthases share the same substrate species.

To date, the three-dimensional structures of three plant monoterpene synthases have been determined by X-ray crystallography, namely, bornyl diphosphate synthase derived from *Salvia officinalis* (Whittington et al., 2002), limonene synthase from *Mentha spicata* (Hyatt et al., 2007), and 1,8-cineol synthase from *S. fruticosa* (Kampranis et al., 2007). Analysis of their structure provided us with information on amino acid residues and related mechanisms of their synthetic reactions; however, the details are still unknown. In order to identify the functions of the domains of geraniol and linalool synthases, we generated chimeric enzymes using a domain swapping approach between these enzymes.

## 2. Results

### 2.1. Identification of the regions responsible for the formation of geraniol (C-1 hydroxylation of geranyl cations)

The amino acid sequences and restriction enzyme recognition sites of geraniol synthase from *P. citriodora* (PcGS) and linalool synthase from *P. hirtella* (PhLS) used in this study are shown in Fig. 3. The transit peptide, which consists of a few dozen residues in the N-terminus that target the initial translation product toward the plastids, was removed from these two monoterpene synthases because there was no significant difference in catalytic activity between full-length and truncated clones of monoterpene synthases from perilla (Bohlmann et al., 1998; Ito and Honda, 2007). The truncated geraniol and linalool synthases share 76% identity at the amino acid level.

To localize the regions of the enzymes responsible for regiospecific hydroxylation, chimeric enzymes (CHs), named CH1, 2, and 3, were constructed (Table I). These

chimeras were made by exchanging corresponding fragments between mutated PCGS (m-PcGS) and mutated PhLS (m-PhLS) restricted by either *AvrII*, *SspI*, or *MfeI*. The sequence from the N-terminus to the *AvrII* restriction site was named domain I, from *AvrII* to *SspI* was named domain II, from *SspI* to *MfeI* was named domain III, and from *MfeI* to the C-terminus was named domain IV (Fig. 4). Domain III includes the DDXD motif, which is highly conservative among monoterpene synthases. CH1, 2, and 3, which have the N-terminus of PhLS and the C-terminus of PcGS and had different proportions of PhLS and PcGS sequences, converted GDP into geraniol. These results suggest that domain IV is important for the formation of geraniol.

Domain IV was further divided into four parts named domains IV-1, -2, -3, and -4 to form CH4–11 (Fig. 4) to identify the essential sequence for geraniol synthesis (Table I). CH4–11 consisted of domains I, II, and III derived from PhLS and a chimeric domain IV. Among CH4–11, CH4 and 5 contained domain IV-4 from PhLS, and the other chimeras harbored domain IV-4 from PcGS. Both CH4 and 5 converted GDP into geraniol and linalool, but the ratio of geraniol were small. However, other chimeras produced geraniol as a principal compound of their reaction products. These results indicate that domain IV-4 of geraniol synthase is indispensable for the formation of geraniol from GDP.

Despite the fact that either CH6, 7, and 8 contained domain IV-4 from PcGS, the percentage of geraniol in reaction products of CH8 was smaller than that in CH6 or 7. This result raises the possibility that domains IV-1 contributes to the formation of geraniol. CH9, 10 and 11 harbored two of the four domains from PhLS in domain IV, but their domain IV-4 were from PcGS (Table I). Geraniol content in reaction products of these chimeras varied, and CH11 produced 100% of geraniol (Table I). These results showed that domains IV-1 and -4 are important for geraniol formation.

## 2.2. The significant regions for the formation of linalool (C-3 hydroxylation of geranyl cations)

Chimeric enzymes of CH13, 14, and 15, either was consisted of N-terminus from PcGS and C-terminus from PhLS. CH13 and 14, which harbored domains III and IV of PhLS, converted GDP into linalool (Table I). Conversely, CH15, which was identical to PhLS only in domain IV, produced geraniol with a trace amounts of linalool. Both

CH13 and 14 contained domain III from PhLS, while CH15 had that from PcGS. Therefore, domain III was further divided into four regions named domains III-a, -b, -c, and -d (Fig. 4) to construct CH16 and 17. CH16, which consisted of a larger percentage of sequence from PcGS in the N-terminus than CH14, produced linalool (Table I). However, CH17, in which domain III-b of CH16 (derived from PhLS) was replaced with the domain from PcGS, formed almost equal amounts of linalool and geraniol (Table I). These results imply that the sequence from domains III-b to C-terminus is important for the formation of linalool.

CH5, 6, 7, and 8 were different in the composition of domain IV-4. CH5 transformed GDP into 93% linalool and 7% geraniol and, CH6, 7, and 8, either had domain IV-4 from PcGS, formed a smaller ratio of linalool (Table I). Therefore, domain IV-4 was important for the formation of linalool, which was consistent with the experimental data that amount of linalool produced in the reactions of chimeras which contained domain IV-4 from PcGS (CH9, 10, 11, etc.) were very little (Table I).

CH18 was arranged with reference to the results of CH5 and 16, i.e., it consisted of domains III-b, -c, and -d and domain IV-4 from PhLS, while the other domains were from PcGS (Table I). Consequently, linalool was detected as a product of CH18. CH19 and 20 were formed for further investigation into the role of domains III-c and -d. CH20, which differed from CH18 only for domain III-d, namely, whether it was from PhLS (CH18) or PcGS (CH20), transformed GDP into geraniol. CH19, which was identical to CH18 except domain III-c, produced linalool, and the percentage of linalool in its reaction products was 94% (Table I). The comparison of reaction products of mutants harboring swapped domains III and IV demonstrated that domains III-b, III-d, and IV-4 were essential for linalool formation.

### 2.3. Roles of domain IV-4 in the regioselective hydroxylation of geranyl cations

We found that the presence of domains IV-1 and -4 from geraniol synthase was necessary for the hydroxylation of the geranyl cation at C-1 (geraniol formation), while the hydroxylation of C-3 (linalool formation) required the presence of domains III-b, III-d, and IV-4 of linalool synthase. Since domain IV-4 was necessary for either geraniol or linalool formation, this region is supposed to be a key area for the regioselective hydroxylation of PcGS and PhLS.

Domain IV-4 was composed of 37 amino acids, and 24 of which were identical in PcGS and PhLS. This suggested that the regiospecificity of additional hydroxyl group was caused by some limited amino acids. Fortunately, CH12, which was made from CH11 by site-directed mutagenesis of H526N and K532Q and deletion of G527, V528, and Q529 (Table I), converted GDP into linalool, as 84% composition of its reaction products (Table I). Namely, the function of geraniol formation of CH11 was changed by mutations of five amino acids. This showed that these five amino acids had very important role in the formation of geraniol.

CH18, which included the essential domains for geraniol formation, except for domain IV-4, and produced linalool, was employed to investigate the role of domain IV-4. In the same manner as for CH12, CH22 was made from CH18 by site-directed mutagenesis of N524H and Q530K and insertion of G525, V526, and Q527 (these five amino acids correspond to the mutated amino acids in CH12). CH22 converted GDP into linalool, but also produced a small amount of geraniol (Table I). As CH18 did not convert GDP into geraniol, these five amino acids might determine the position of the additional hydroxyl group. However, more mutants were constructed by site-directed mutagenesis, but they failed to specify amino acids that are indispensably control the regiospecificity of linalool synthase.

For all of the enzymatic reactions performed in this study, the only monoterpene compounds that were detected in the reaction products were geraniol and linalool.

### 3. Discussion

#### 3.1. Regiospecificity-determinants were not identical between geraniol and linalool synthases

Both geraniol and linalool synthases utilize GDP as substrate, but the position of the additional hydroxyl group differs between geraniol (C-1) and linalool (C-3) (Fig. 2). The deduced amino acid sequences of PcGS and PhLS showed 76% identity for the full-length proteins, and 69% for domain I, which is immediately downstream of the RRX<sub>8</sub>W motif, and 65% for domain IV-4, which is just before the stop codon (Fig. 4). The low amino acid identity of these domains could indicate that they are more pertinent to regioselective hydroxylation than the other parts of the enzymes (Fig. 2).



Consequently, the domain swapping experiment with PcGS and PhLS revealed that domain IV-4 is the most indispensable region for both linalool and geraniol syntheses.

The chimeric enzymes CH11 and CH19 were based on the sequences of linalool and geraniol synthases in which the required regions were replaced to produce geraniol and linalool, respectively. CH11, which produced geraniol, is identical to PhLS, except for domains IV-1 and -4, and CH11 and PhLS shared 95% identity at the amino acid level. CH19, which harbors domains III-b, III-d and IV-4 from PhLS and is capable of linalool formation, shared 94% identity with PcGS at the amino acid level. Two putative pathways of linalool formation have been considered, one was direct hydroxylation of geranyl cation and the other was pathway via LDP (Degenhardt et al., 2009). The fact that the regiospecificity determining regions for geraniol and linalool synthases were not identical may indicate that the reaction mechanisms are different between these two synthases.

Terpene synthases share similar three-dimensional structures despite the absence of significant sequence similarity (Degenhardt et al., 2009). Geraniol and linalool synthases share the highest similarity with bornyl diphosphate synthase among the three monoterpene synthases whose three-dimensional structures have been elucidated by X-ray crystallography. Their sequence similarity with bornyl diphosphate synthase was also the greatest when the amino acids of domains III-b and -d and IV-1 and -4, which were shown to be important for the regiospecificity of the additional hydroxyl group, were compared. On the basis of the previously elucidated structure of bornyl diphosphate synthase from *S. officinalis* (Protein Data Bank ID code 1N1B), three-dimensional models of CH11 and CH19 were created using PyMOL and SWISS-MODEL (Fig. 5). The regions shown to be indispensable for regioselectivity in the domain-swapping studies were found to be in the surrounding area of the active site cavity of the enzyme; domains III-b and -d and IV-4 surrounded the cavity in CH19, while domains IV-1 and 4 were located on one surface of the cavity in CH11 (Fig. 5). In other words, the same circumstances could be illustrated from the standpoint of structural elements like alpha helices, CH11 consisted of the helices G2–H2 and K derived from PcGS and the other helices derived from PhLS while CH19 harbored helices D, F, G1 and K from PhLS and the other helices from PcGS (Fig. 5)(Whittington et al., 2002). This difference in the arrangement of the important regions



between CH11 and CH19 is considered to affect the configuration of carbocation intermediates and their electron localization.

### 3.2. Five amino acids in domain IV-4 were indispensable for the formation of geraniol

Since CH11 and CH19 indicated that domain IV-4 was an indispensable region for geraniol and linalool syntheses, further investigations on domain IV-4 were performed; CH12 was constructed based on CH11 by transforming five residues in domain IV-4. As a consequence, the product specificity of the chimeric enzyme was completely changed, and the percentage of geraniol was only 14% in its reaction products. This indicates that those five amino acids in domain IV-4 are essential for the determination of the position of the additional hydroxyl group during the formation of geraniol. These five amino acids were H526, G527, V528, and Q529, which are located in the J-K loop, and K532, which is in helix K. A three-dimensional model of CH12 was constructed and compared with that of CH11. It was found that the J-K loop of CH11 was longer than that of CH12 by three residues (G527, V528, and Q529) and was prominent toward the active site cavity with GDP (Fig. 6). Due to the difference in the length of the J-K loop, the spatial arrangement of the amino acid at position 526 was altered dramatically in CH12 from CH11. The model indicated that the residue at position 526 corresponds to F578 in bornyl diphosphate synthase from *S. officinalis*, and the side chain of F578 seemed to contribute to the binding of the substrate GDP and stabilized the carbocation intermediate (Whittington et al., 2002). For CH11, the imidazole ring of H526 protruded toward the substrate in the active site cavity, which seemed to interact with GDP to stabilize the carbocation intermediate (Fig. 6). For CH12, the side chain of N526 extended toward the inverse side of the active site cavity, suggesting that the interaction between N526 and the ligand is significantly weaker than in CH11 (Fig. 6). The distance between H526 of CH11 and D299 which was located at the N-terminal end of the DDXXD motif and was 8.4Å, while that between N526 of CH12 and D299 was 16.5Å, when they were analyzed by PyMOL. These findings indicate that the contribution of residues in the loop to fix and localize the positive charge of the carbocation intermediate leads to the formation of geraniol in the reaction of CH11, whereas a more stable tertiary allyl carbocation intermediate was formed without or with weak interactions between the amino acids and the ligand for CH12, which

produced linalool. The difference in these five amino acids between PcGS and PhLS is conserved among the geraniol and linalool synthases previously cloned from cultivated and wild perilla species (Ito and Honda, 2007; Masumoto et al., 2010). It seemed that the regioselectivity of the hydroxyl group between geraniol and linalool synthases depends on the distance between the side chain of the amino acid (position N526 in PhLS and position H524 in PcGS) and the ligand. It is expected that X-ray crystal structure analysis of the chimeric enzymes constructed in this study as well as geraniol and linalool synthases might confirm our results and produce more findings about acyclic monoterpene synthases, and efforts to determine their X-ray crystal structures are in progress.

### 3.3 All chimeric enzymes created in this study formed only geraniol and/or linalool

Almost all of the chimeric enzymes generated in this study, except for one (CH10), were able to transform GDP (Table I), while almost half of the chimeric enzymes were inactive in previous domain swapping studies on monoterpene synthases (El Tamer et al., 2003; Katoh et al., 2004). The results of this study were also very different from our previous study in which almost all chimeras made from a combination of d-limonene synthase from *P. frutescens* (GenBank accession no. BAA08367) and l-limonene synthase from *Schizonepeta tenuifolia* (GenBank accession no. AAG01140) were inactive (data not shown) (Maruyama, 2002). Interestingly, all chimeras formed only geraniol and/or linalool, whereas previous studies reported that chimeric enzymes simultaneously transformed GDP into multiple types of monoterpene compounds (El Tamer et al., 2003; Katoh et al., 2004; Peters and Croteau 2003). In this respect, the chimeras produced in this study, as well as their parental enzymes, are unique. Previous studies of multifunctional chimeras used monoterpene cyclases as the parental enzymes, while the geraniol and linalool synthases used in this study are members of a group of enzymes that produce acyclic terpenes. It is known that both cyclic and acyclic monoterpene synthases utilize GDP as a substrate, and their three-dimensional structures are similar to each other. However, linalool and geraniol synthases dephosphorylate GDP to form a geranyl cation and add a hydroxyl group to the intermediate without modifying its carbon skeleton (Degenhardt et al. 2009; Iijima et al., 2004; Jia et al., 1999). While, cyclases also form a geranyl cation by dephosphorylating

GDP, they further convert the intermediate into an  $\alpha$ -terpinyl cation by cyclization and produce various types of monoterpene compounds. In this case, subtle differences in the folding pattern and electron localization of the  $\alpha$ -terpinyl cation (Degenhardt et al. 2009) could create a variety of products, which may be followed by the formation of more diverse by-products than by geraniol and linalool synthases.

#### 4. Conclusions

The mechanisms underlying the regiospecific hydroxylation of geraniol and linalool synthases were investigated by a domain swapping approach. It was shown that important regions for the formation of geraniol and linalool synthases were not identical, namely, domains IV-1 and -4 for geraniol, and domains III-b, III-d and IV-4 for linalool, respectively. Further, five amino acids in domain IV-4 were indispensable for the formation of geraniol. According to the three-dimensional structural models of the synthases, these five residues seemed to be responsible for the different spatial arrangements of the amino acid at H524 in geraniol synthase (N526 is the corresponding residue in linalool synthase). It was suggested that the side-chains of these five amino acids in combination with the relevant domains localized the positive charge of the carbocation intermediate to determine the position for introducing the hydroxyl group to the substrate.

#### 5. Experimental

##### 5.1. General experimental procedures

Authentic geraniol and linalool and all chemical reagents and solvents were purchased from Nacalai Tesque, Inc. or Wako Pure Chemical Industries Co. Gel and plasmid extraction kits were purchased from Takara. The His-tagged protein purification kit was from Qiagen, and vectors and competent *Escherichia coli* cells were from Invitrogen. Solid Phase Microextraction (SPME) fiber (100  $\mu$ m Polydimethylsiloxane) was from Supelco. Polymerase chain reactions were performed on a Thermal Cycler Personal (Takara). GC-MS measurements were carried out on a G-9000M gas chromatograph (HITACHI) connected to an M-7000 (HITACHI). GC-FID measurements were

performed on a G-5000 (HITACHI). DNA sequencing was performed by Bio Matrix Research, Inc.

## 5.2. Plant materials

All perilla plants used in this study were grown in the Experimental Station for Medicinal Plant Research, Graduate School of Pharmaceutical Sciences, Kyoto University. They have been bred and kept as pure lines as reported previously (Ito and Honda, 2007). The strain numbers, oil types, and species used in this study were as follows: *P. citriodora* [Strain No. 87, C type] and *P. hirtella* [Strain No. 5073, PK type].

## 5.3. DNA clones from perilla and construction of chimeric enzymes

cDNAs encoding geraniol synthase from *P. citriodora* (GenBank accession no. DQ088667) named PcGS and linalool synthase from *P. hirtella* (GenBank accession no. FJ644546) named PhLS, which were both cloned previously (Ito and Honda, 2007; Masumoto et al., 2010), were used as the parental enzymes of the chimeric enzymes. The cDNAs were truncated by 177 or 183 bases at the N-terminus of PcGS or PhLS, respectively, and ligated to the pCR T7/CT-TOPO vector (Invitrogen). To create restriction enzyme recognition sites at the same relative position between the two sequences, a site-directed mutagenesis approach was adopted using a QuikChange II Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's protocol. PcGS was mutated as follows to create m-PcGS: deletion of the *SspI* restriction site (I90: ATT→ATC), creation of *AvrII* (A103 and L104: GCT→GCC and CTT→CTA, respectively), *AfeI* (I271: ATT→ATA), *AflIII* (L350 and K351: CTC→CTT and AAA→AAG, respectively), *SpeI* (T445 and S466: ACA→ACT and TCA→AGT, respectively), *HindIII* (E483 and A484: GAG→GAA and GCG→GCT, respectively), and *PvuII* (A506 and A507: GCC→GCA and GCA→GCT, respectively) restriction sites. PhLS was also mutated to create m-PhLS: *AvrII* (A102 and L103: GCT→GCC and CTT→CTA, respectively), *AfeI* (I273 and A274: ATT→ATA and GCT→GCG, respectively), *PstI* (L311 and Q312: CTA→CTG and CAA→CAG, respectively), *AflIII* (L352: CTC→CTT), *MfeI* (I400: ATA→ATT), *SpeI* (T447 and S448: ACA→ACT and TCT→AGT, respectively), *HindIII* (E485 and A486: GAG→GAA and GCG→GCT, respectively), and *PvuII* (A507 and A508: GCG→GCA and GCA→GCT, respectively).

However, both PcGS and PhLS were mutated to m-PcGS and m-PhLS without changing the integrity of the encoded amino acid sequences.

The chimeric genes were constructed as follows: m-PcGS and m-PhLS were digested with the appropriate restriction enzyme or combination of enzymes at the sites shown in Fig. 3. The digested products were electrophoresed, and fragments of the correct size were isolated from the agarose gel using NucleoSpin<sup>®</sup> Extract II (Macherey-Nagel) and ligated to create the chimeric cDNAs shown in Table I. All chimeric genes constructed were sequenced to confirm the correct sequences.

#### 5.4. Heterologous expression of cDNAs in *E. coli* and His-tagged protein purification

Each plasmid containing the parental or chimeric cDNA was introduced into *E. coli* BL21 (DE3) pLysS cells (Invitrogen) for heterologous expression. Single colonies were isolated, grown for 6–8 h at 37 °C with shaking in LB liquid medium, and then inoculated into 40 mL TB liquid medium containing 0.1 mg·ml<sup>-1</sup> ampicillin. After 1.5–2 h shaking at 37 °C, isopropyl thio-β-galactoside was added to a final concentration of 1 mM and the culture was transferred to 16 °C and cultured for 26 h with shaking (220 rpm). The cultured cells were harvested by centrifugation, resuspended in a Tris buffer, sonicated, and applied to a Ni-NTA Spin-column (Qiagen) to isolate the His-tagged recombinant proteins, as described previously (Kumeta and Ito, 2010). Protein concentrations were determined by the Bradford assay (Bio-Rad) using bovine serum albumin as the standard.

#### 5.5. Enzymatic assays and GC-FID/GC-MS analysis

Enzymatic assays: the reaction mixtures were composed of 50 mM Tris buffer (pH 7.0), 10% (v/w) glycerol, 1 mM dithiothreitol, 0.5 mM MnCl<sub>2</sub>, 20 mM MgCl<sub>2</sub>, and 28 μM GPP. Purified recombinant proteins were added to the mixtures to a final concentration of 10% (v/v). Completed mixtures were kept at 30 °C for 14 h for the enzymatic reaction. The enzymatic assay for each chimera was performed in duplicate. In some cases, the amount of reaction products was very small for short-time reactions and they were not detectable with GC analyses. Therefore, a reaction time of 14 h was used for all experiments.

SPME-GC-MS analysis: the enzymatic reaction was performed in a solution of 250

μL in a glass vial (4 mL) sealed with a polytetrafluoroethylene-coated silicone rubber septum. A preparation devoid of cofactors ( $Mn^{2+}$  and  $Mg^{2+}$ ) and extracts prepared from *E. coli* harboring a plasmid ligated with the coding region of PhLS in the reverse direction were employed as controls for the assays. After the enzymatic reaction, SPME fiber (100 μm Polydimethylsiloxane; Supelco) was inserted into the headspace of the vial at room temperature for 30 min and then transferred to the injector of a GC-MS (G-9000M gas chromatograph connected to an M-7000 (HITACHI)). The SPME fiber was heated for the desorption of compounds for 10 min at 200 °C. The compounds were separated on a TC-WAX column (60 m × 0.25 mm, 0.25 μm film thickness; GL Sciences, Inc.) with a temperature program starting at 100 °C, holding at this temperature for 5 min, then increasing by 5 °C·min<sup>-1</sup> to 200 °C, and holding at this temperature for 5 min. Helium was used as the carrier gas, and the column flow rate was set to 1.0 mL·min<sup>-1</sup>. The eluted compounds were identified by comparison of their ion spectra and relative retention times of the authentic standards.

GC-FID analysis: the enzymatic reaction mixtures (1 mL) were prepared in screw-capped glass tubes and overlaid with 2 mL pentane. After incubation, the reaction mixtures were extracted three times with pentane, while 12 nmol limonene in 10 μL hexane were added to each tube as an internal standard just before extraction. The pentane fractions were combined and concentrated under a N<sub>2</sub> gas stream to be injected into a GC system equipped with the same column as for SPME-GC-MS. The operation conditions were as follows: injector: 160 °C; FID: 210 °C; oven program starting at 50 °C, holding at this temperature for 5 min, then increasing by 3 °C·min<sup>-1</sup> to 200 °C, and holding at this temperature for 5 min. The amount of reaction products was calculated based on GC-FID analysis. The percentage of each compound in the reaction products was shown in Table I.

### 5.6. Molecular modeling

PyMOL and SWISS-MODEL (<http://swissmodel.expasy.org/>) were used to develop the three-dimensional models of CH11, 12 and 19 based on the structure of bornyl diphosphate synthase from *Salvia officinalis* (Protein Data Bank ID code 1N1B).

## Acknowledgement

This work was partially supported by a Grant-in-Aid for JSPS Fellows (No. 12J08932).



## Figure legends

### Figure 1

Putative biosynthetic pathways of the oil constituents of perilla. (1) Geranyl diphosphate; (2) linalool; (3) geraniol; (4) limonene; (5) citral; (6) perillaldehyde; (7) piperitenone; (8) perillene; (9) perillaketone; (10) elsholtzia ketone; and (11) shisofuran.

### Figure 2

Putative reaction mechanisms for the conversion of GDP to geraniol and linalool by monoterpene synthases. OPP denotes the diphosphate moiety.

### Figure 3

Alignment of the amino acid sequences of geraniol synthase (PcGS) and linalool synthase (PhLS). The borders of the domains of the monoterpene synthases used in this study are indicated by arrows. The amino acid residues marked with an asterisk were mutated in the site-directed mutagenesis study of domain IV-4. The closed circles show the RRX<sub>8</sub>W motif and the thick bar indicates the DDXXD motif.

### Figure 4

Primary structure of geraniol and linalool synthases cloned into the pCRT7/CT-TOPO vector and divided into domains I–IV. Domain III was further divided into four domains, domains III-a–d, and domain IV is divided into domains IV-1–4. The restriction enzymes used in this study and the amino acid positions bordering the domains are shown. The similarity of the two synthases is shown as percentage of amino acid identity for each domain. The DDXXD motif is included in domain III-b.

### Figure 5

Three-dimensional homology models of CH19 (A) and CH11 (B) constructed with reference to the structure of bornyl diphosphate synthase from *S. officinalis*. Capital letters are the codes of alpha helices.

(A) The regions required for linalool formation (yellow parts) surround the active site cavity of the enzyme. The helices and loops colored yellow denote the regions required

for linalool formation, and the red parts indicate the DDXXD motif.

(B) The domains required for geraniol formation (green parts) were located on one surface of the cavity of CH11. The helices and loops colored green denote the regions required for the formation of geraniol from GDP, and the red parts indicate the DDXXD motif.

## Figure 6

C-terminal domain (amino acids between position 247 and the C-terminus) in the three-dimensional homology models of CH11 (A) and CH12 (B) based on the structure of bornyl diphosphate synthase from *S. officinalis*. The large cavity in (A) and (B) is considered to be the active site of the enzyme.

(A) The imidazole group of H526 in CH11 extends close to the substrate in the active site cavity.

(B) The side chain of N526 in CH12 seems to be located on the opposite side of the active site cavity.

Green: five amino acids in the J-K loop that are important for geraniol formation; yellow: the five mutated amino acids introduced into CH11; red: the DDXXD motif.

**Table I**

Schematic representation of the monoterpene synthase chimeras and their products and ratio of each compound in reaction products

| Clone name        | Domain                |                        |                        |                        | Percentage of<br>linalool or geraniol<br>in reaction products |          |
|-------------------|-----------------------|------------------------|------------------------|------------------------|---|----------|
|                   | I (69% <sup>a</sup> ) | II (83% <sup>a</sup> ) | III(79% <sup>a</sup> ) | IV (71% <sup>a</sup> ) | Linalool  | Geraniol |
| m-PhLS            |                       |                        |                        |                        | 100   | 0        |
| m-PcGS            |                       |                        |                        |                        | 0   | 100      |
| CH1               |                       |                        |                        |                        | 0   | 100      |
| CH2               |                       |                        |                        |                        | 0   | 100      |
| CH3               |                       |                        |                        |                        | 11  | 89       |
| CH4               |                       |                        |                        | 2                      | 93  | 7        |
| CH5               |                       |                        |                        | 4                      | 93  | 7        |
| CH6               |                       |                        |                        | 3                      | 10  | 90       |
| CH7               |                       |                        |                        | 2                      | tr.   | 100      |
| CH8               |                       |                        |                        | 1                      | 23  | 77       |
| CH9               |                       |                        |                        | 3                      | 21  | 79       |
| CH10              |                       |                        |                        | 2 4                    | tr.   | tr.      |
| CH11              |                       |                        |                        | 1 4                    | tr.   | 100      |
| CH12 <sup>b</sup> |                       |                        |                        | 1                      | 86  | 14       |
| CH13              |                       |                        |                        |                        | 100   | 0        |
| CH14              |                       |                        |                        |                        | 100   | 0        |
| CH15              |                       |                        |                        |                        | tr.   | 100      |
| CH16              |                       |                        | a                      |                        | 100   | 0        |
| CH17              |                       |                        | b                      |                        | 44  | 56       |
| CH18              |                       |                        | a                      | 4                      | 100   | 0        |

|                   | Domain |  |   |  |   |   |  |  |   | Percentage of<br>linalool or geraniol<br>in reaction products |     |
|-------------------|--------|--|---|--|---|---|--|--|---|---|-----|
| CH19              |        |  | a |  | c |   |  |  |   | 94  | 6   |
| CH20              |        |  | a |  |   | d |  |  | 4 | 0   | 100 |
| CH21 <sup>c</sup> |        |  | a |  |   |   |  |  |   | 85  | 15  |
| CH22 <sup>d</sup> |        |  | a |  |   |   |  |  |   | 81  | 19  |

Chimeric cDNAs are represented by gray and white boxes indicating the PcGS (gray) and PhLS (white) peptide regions. Domains III-a–d and domains IV-1–4 are represented as “a–d” and “1–4,” respectively.

<sup>a</sup>The similarity of geraniol and linalool synthases is shown as percentage of amino acid identity for each domain.

<sup>b</sup>CH12 was constructed by the replacement of H526N and K532Q and deletion of G527, V528, and Q529 in CH11.

<sup>c</sup>CH21 was identical to CH18, except for G525, V526, and Q527.

<sup>d</sup>CH22 was constructed by the mutation of CH21 at N524H and Q530K.

–: not detected.

tr.: trace.

## References

- Bohlmann, J., Meyer-Gauen, G., Croteau, R., 1998. Plant terpenoid synthases: molecular biology and phylogenetic analysis. *Proc. Natl. Acad. Sci. USA.* 95, 4126-4133.
- Chappell, J., 1995. The biochemistry and molecular biology of isoprenoid metabolism. *Plant Physiol.* 107, 1-6.
- Chen, F., Tholl, D., D'Auria, J.C., Farooq, A., Pichersky, E., Gershenzon, J., 2003. Biosynthesis and emission of terpenoid volatiles from *Arabidopsis* flowers. *The Plant Cell.* 15, 481-494.
- Chung, B.H., Lee, H.Y., Lee, J.S., Young, C.Y., 2006. Perillyl alcohol inhibits the expression and function of the androgen receptor in human prostate cancer cells. *Cancer Lett.* 236, 222-228.
- Degenhardt, J., Köllner, T.G., Gershenzon, J., 2009. Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry.* 70, 1621-1637.
- E.I., Tamer, M.K., Lückner, J., Bosch, D., Verhoeven, H.A., Verstappen, F.W.A., Schwab, W., van Tunen, A.J., Voragen, A.G.J., de Maagd, R.A., Bouwmeester, H.J., 2003. Domain swapping of Citrus limon monoterpene synthases: impact on enzymatic activity and product specificity. *Arch. Biochem. Biophys.* 411, 196–203.
- Honda, G., Koga, K., Koezuka, Y., Tabata, M., 1984. Antidermatophytic compounds of *Perilla frutescens* BRITTON var. *crispa* DECNE. *Shoyakugaku Zasshi.* 38, 127-130.
- Honda, G., 1996. Genetic control of the essential oil components in *Perilla*. *FFI J.* 169, 9-15.
- Iijima, Y., Gang, D.R., Fridman, E., Lewinsohn, E., Pichersky, E., 2004.

Characterization of Geraniol Synthase from the Peltate Glands of Sweet Basil. *Plant Physiol.* 134, 370-379.

Hyatt, D.C., Youn, B., Zhao, Y., Santhamma, B., Coates, R.M., Croteau, R.B., Kang, ChulHee, 2007. Structure of limonene synthase, a simple model for terpenoid cyclase catalysis. *PNAS.* 104, 5360-5365.

Ito, M., Toyoda, M., Honda, G., 1999. Chemical composition of the essential oil of *Perilla frutescens*. *Nat. Med.* 53, 32-36.

Ito, M., Honda, G., 2007. Geraniol synthases from perilla and their taxonomical significance. *Phytochemistry.* 68, 446-453.

Jia, J.W., Crock, J., Lu, S., Croteau, R., Chen, X.Y., 1999. (3R)-Linalool Synthase from *Artemisia annua* L.: cDNA Isolation, Characterization, and Wound Induction. *Arch. Biochem. Biophys.* 372, 143-149.

Kampranis, S.C., Ioannidis, D., Purvis, A., Mahrez, E., Ninga, E., Katerelos, N.A., Anssour, S., Dunwell, J.M., Degenhardt, J., Makris, A.M., Goodenough, P.W., Johnson, C.B., 2007. Rational conversion of substrate and product specificity in a *Salvia* monoterpene synthase: structural insights into the evolution of terpene synthase function. *The Plant Cell.* 19, 1994-2005.

Katoh, S., Hyatt, D., Croteau, R., 2004. Altering product outcome in *Abies grandis* (-)-limonene synthase and (-)-limonene/(-)- $\alpha$ -pinene synthase by domain swapping and directed mutagenesis. *Arch. Biochem. Biophys.* 425, 65-76.

Kumeta, Y., Ito, M., 2010. Characterization of d-guaiene synthases from cultured cells of *aquilaria*, responsible for the formation of the sesquiterpenes in agarwood. *Plant Physiol.* 154, 1998-2007.

Maruyama, T., 2002. Studies on stereoselectivity in limonene biosynthesis. PhD thesis.

Kyoto university, Kyoto

Masumoto, N., Korin, M., Ito, M., 2010. Geraniol and linalool synthases from wild species of perilla. *Phytochemistry*. 71, 1068-1075.

Peters R. J., Croteau R. B., 2003. Alternative termination chemistries utilized by monoterpene cyclases: chimeric analysis of bornyl diphosphate, 1,8-cineole, and sabinene synthses. *Arch. Biochem. Bioshys.* 417, 203-211.

Pichersky, E., Lewinsohn, E., Croteau, R., 1995. Purification and characterization of S-linalool synthase, an enzyme involved in the production of floral scent in *Clarkia breweri*. *Arch. Biochem. Biophys.* 316, 803-807.

Whittington, D.A., Wise, M.L., Urbansky, M., Coates, R.M., Croteau, R.B., Christianson, D.W., 2002. Bornyl diphosphate synthase: structure and strategy for carbocation manipulation by a terpenoid cyclase. *PNAS*. 99, 15375-15380.

Yamasaki, Y., Kunoh, H., Yamamoto, H., Akimitsu, K., 2007. Biological roles of monoterpene volatiles derived from rough lemon (*Citrus jambhiri* Lush) in citrus defense. *J. Gen. Plant Pathol.* 73, 168-179.

Yuba, A., Yazaki, K., Tabata, M., Honda, G., Croteau, R., 1996. cDNA cloning, characterization, and functional expression of 4S-(-)-Limonene synthase from *Perilla frutescens*. *Arch. Biochem. Biophys.* 332, 280-287.



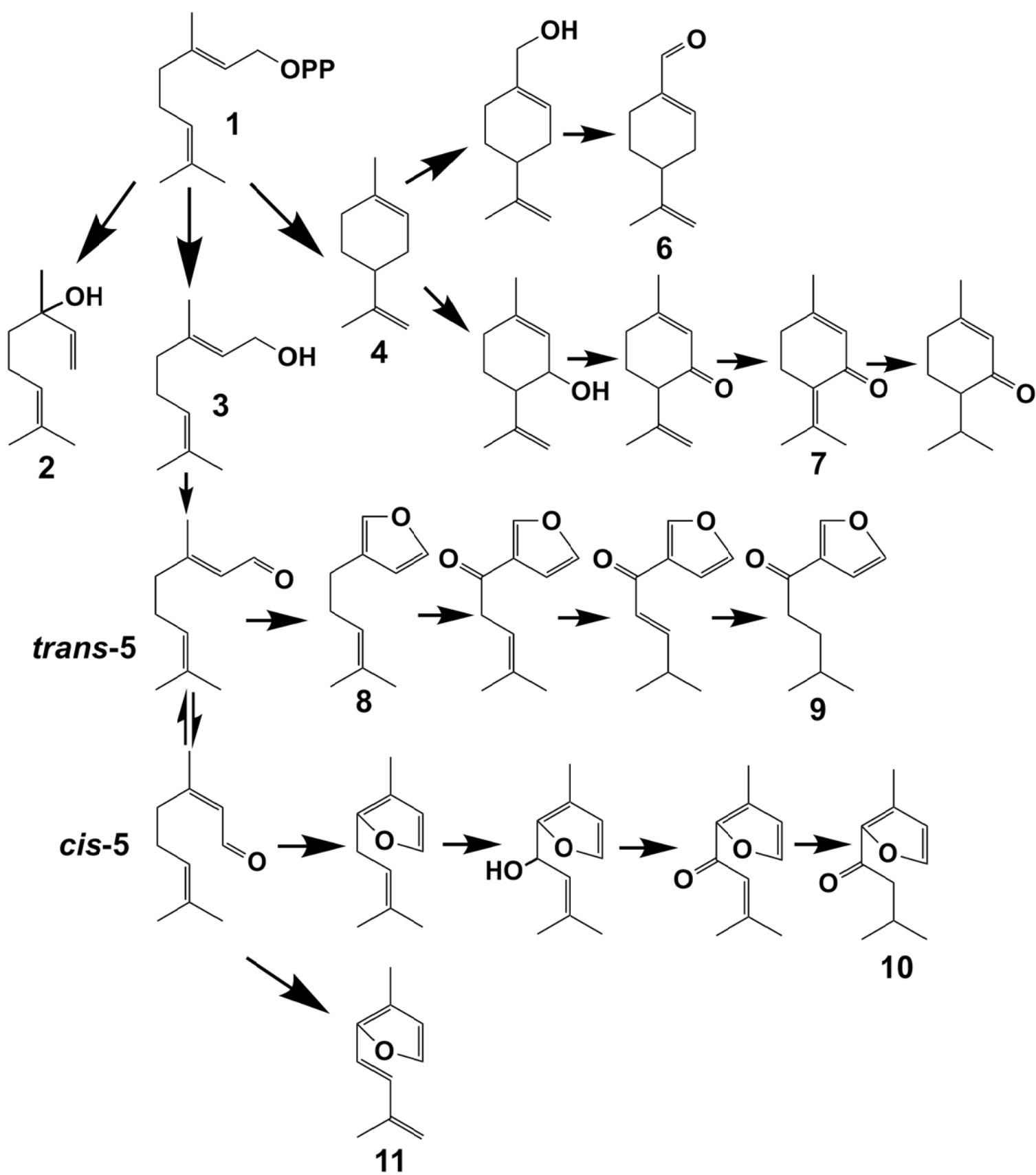


Figure 1

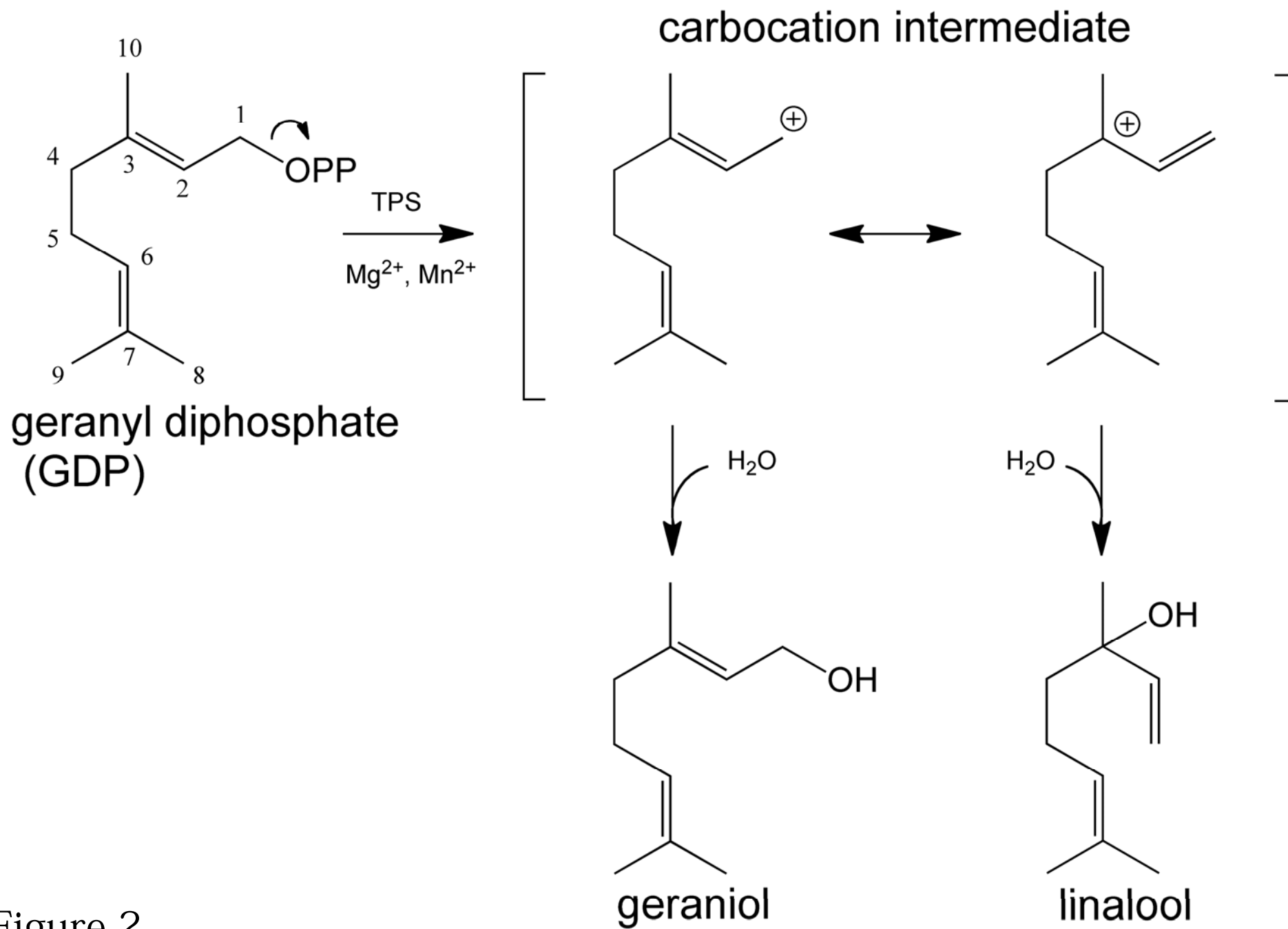


Figure 2

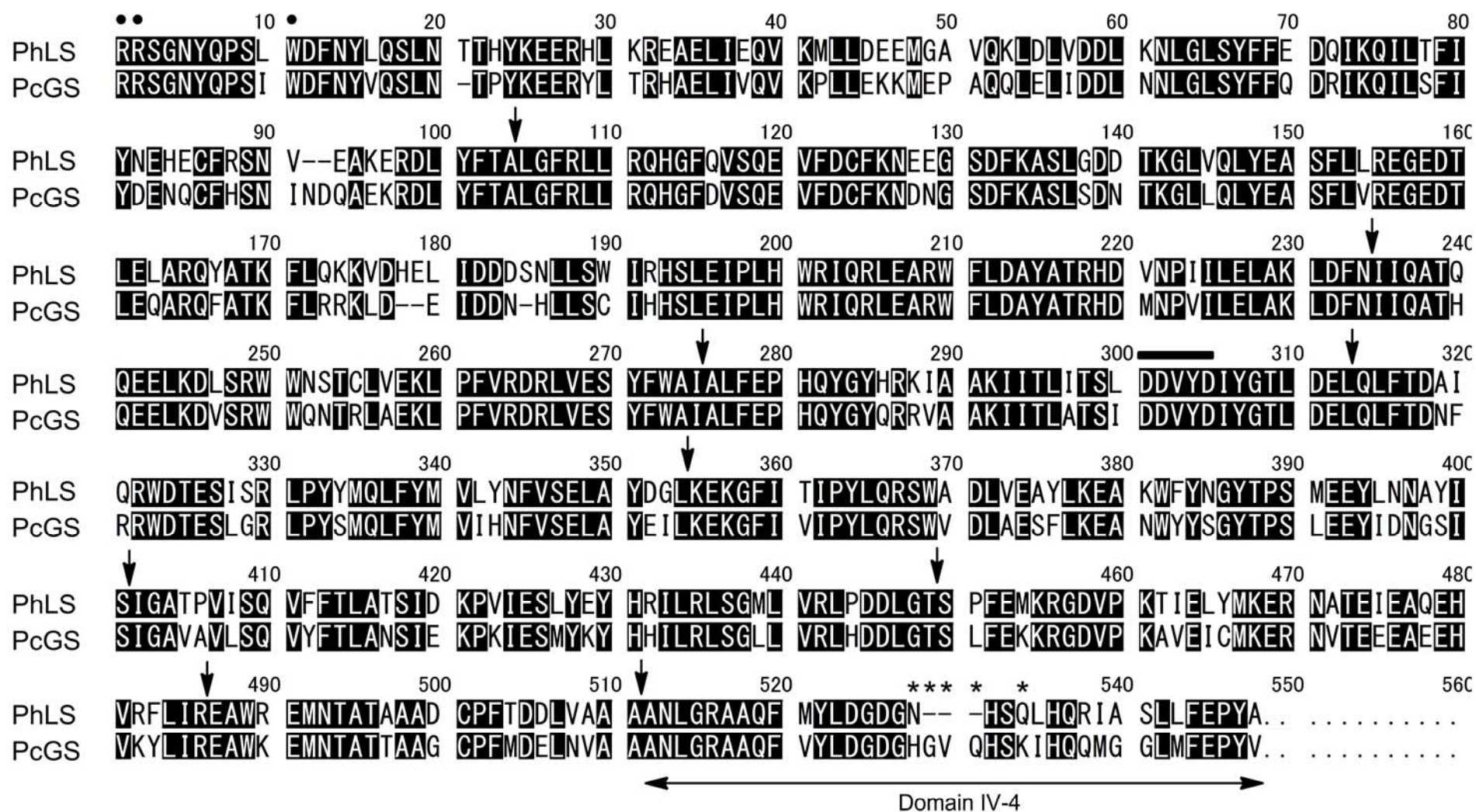


Figure 3

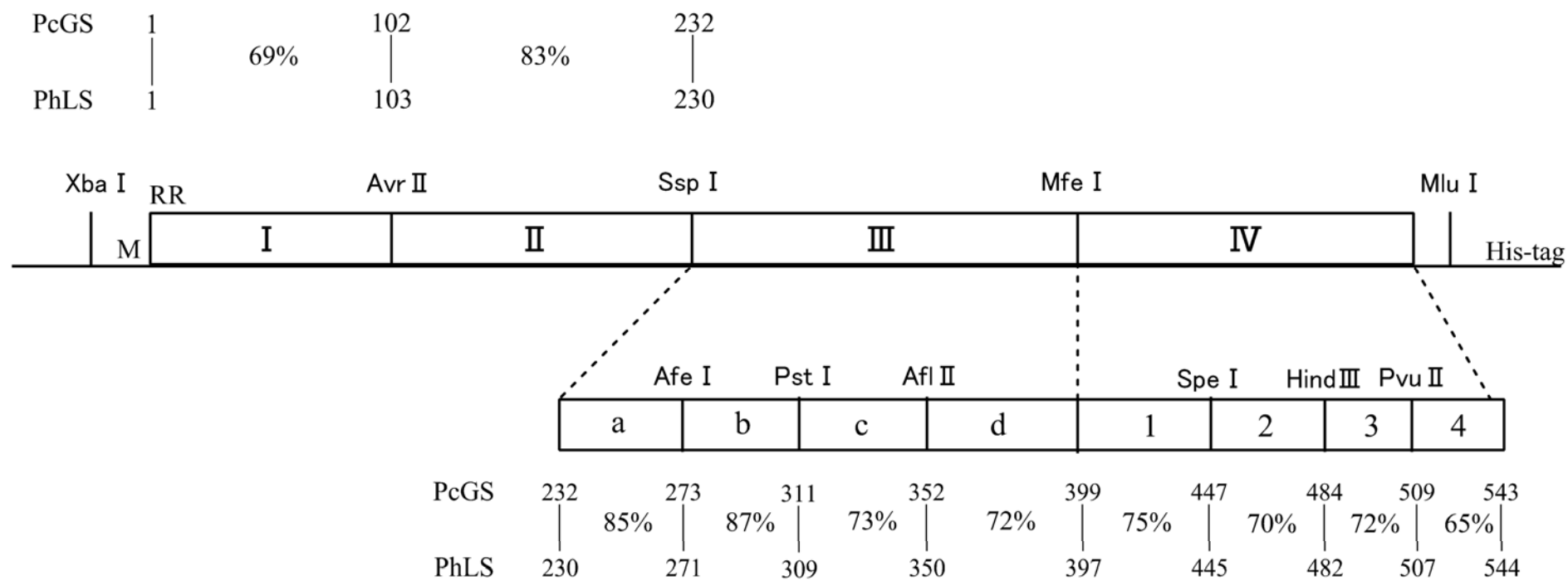


Figure 4

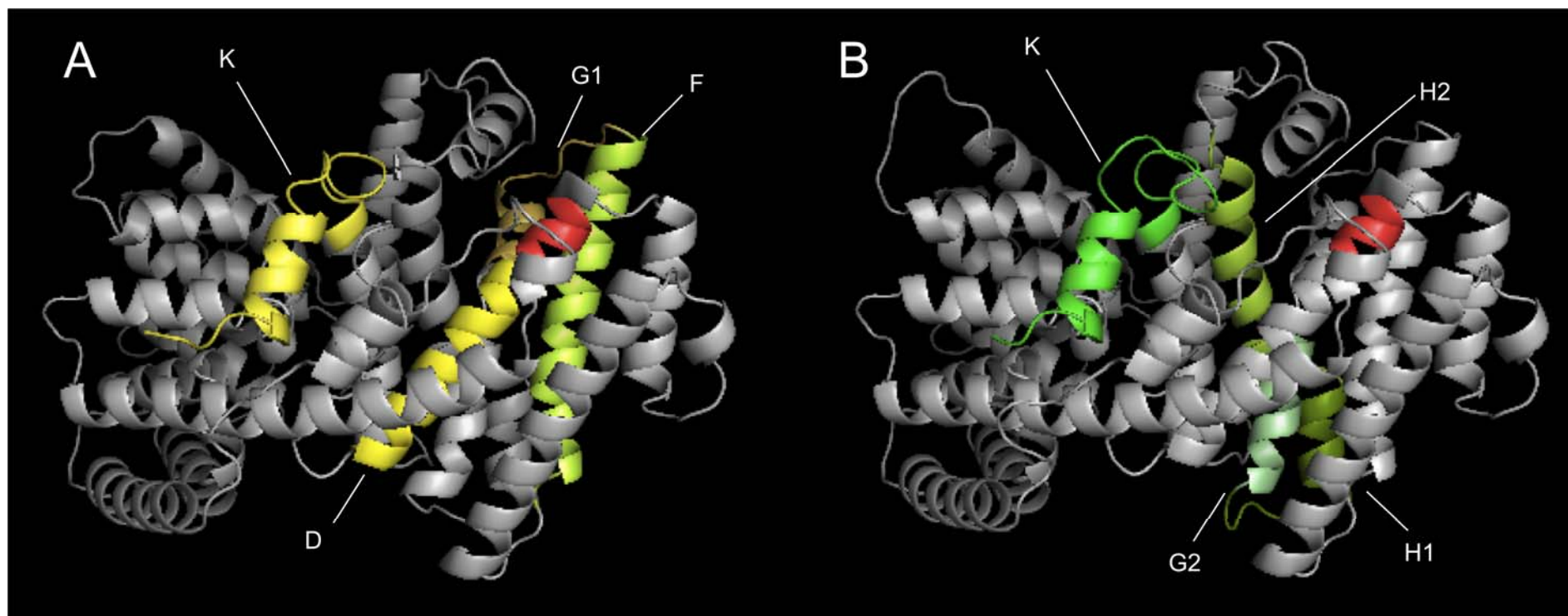
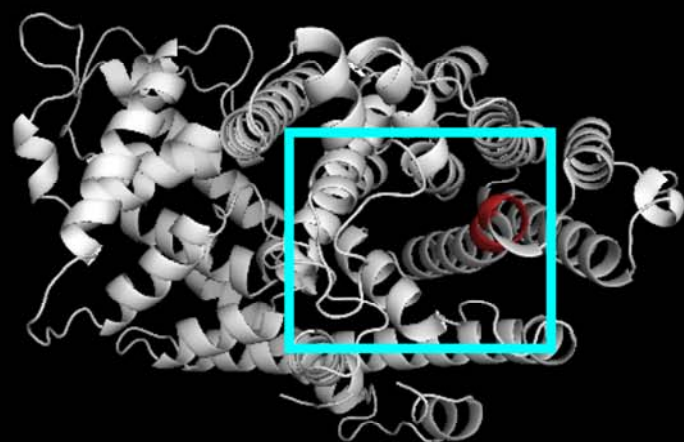
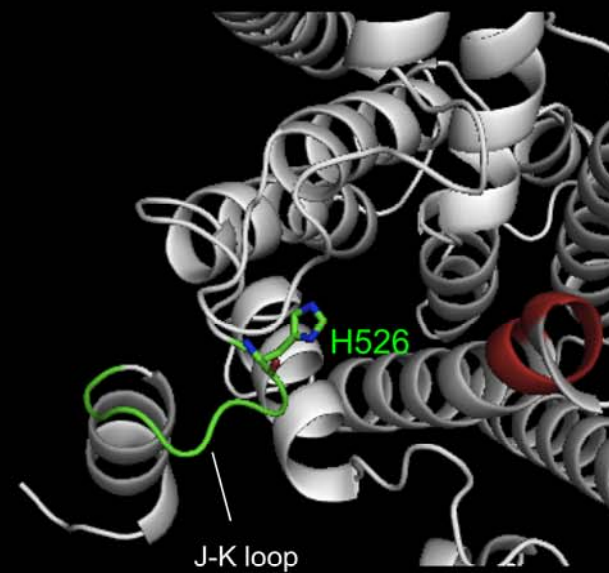


Figure 5



A



B

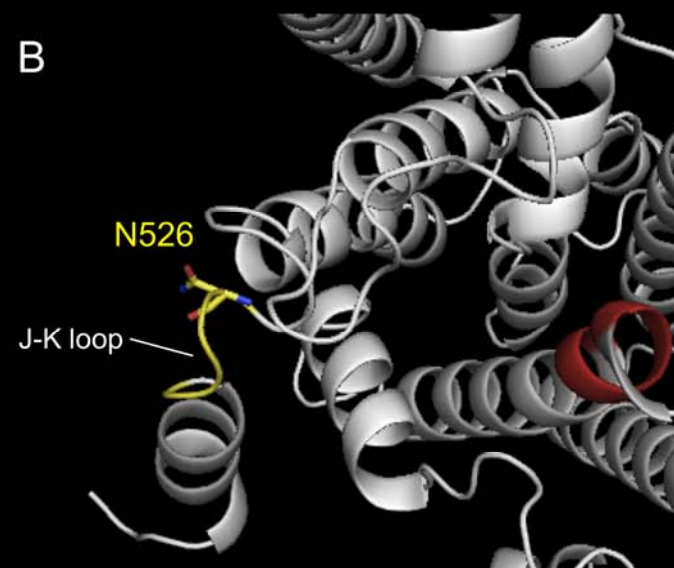


Figure 6